

# Real Time Adaptive Tracking System Using Computer Vision

**Ayonika Paul**

*P.G. Student*

*Electronics Engineering, Pillai College of Engineering,  
New Panvel, India*

*University of Mumbai*

**Email:** *ayonika\_paul@yahoo.com*

## **Abstract**

*This project studies long-time object tracking in a sequence of frames. In this project, a detector is trained with specimens found on the path of a tracker that itself does not rely on the object detector. We attain high robustness and outdo current adaptive tracking-by-detection (11) approaches by decoupling object tracking and object detection. A substantial reduction of calculating time is attained by means of simple features for object detection and by using a cascaded method. The object location is marked in each frame. The task is to find the position of object in that frame else it must notify that the object is absent in the consecutive frames. We have developed a Real Time Tracking framework. The task of long-time tracking is divided as follows: Tracking, Learning and Detection. The tracker must follow the marked object of interest in consecutive frames. The detector restricts all observed appearances and amends the tracker when required. To evade these blunders there forth, the learning approximates the detector's blunders and re-evaluates it. This project studies methods to recognize the detector's faults and learn from, by developing a learning method with the help of "experts" which will estimate these blunders. We call it the P-N learning. With the help of RAT and P-N learning, our real-time processing can be described as an extremely integrated arrangement providing very precise object detection with RGB-D sensor.*

**Keywords:** *Real time object tracking, RAT framework, Judge and track objects from dissimilar distances, Learning, Experts.*

## **INTRODUCTION**

Researchers in computer vision develop mathematical techniques based on how the human visual system interprets environmental stimuli. It is to extract information about physical subjects built on camera images. Computer vision approaches are applied to quality inspection, scene reconstruction, robot management, optical character recognition and object organisation. A field of study in computer vision where approaches are studied that evaluate the position of object of interest in successive video frames is object tracking. Due to extensive use of advanced computers and due to the availability of high-tech and budget friendly video cameras, the necessity for

automatic video examination have influenced applying object tracking algorithms in automatic annotation of video data, vehicle navigation, automated surveillance, human-computer communication and traffic monitoring. Consider a video stream where several objects move in and out of the camera frame. The goal of this project is to identify the object's bounding box automatically. The box must define the object of interest or notify that the object is not present in following frames. In order to process the video stream frame to frame, the process must run indefinitely long which we call as long-term tracking. The long-term tracking has a few drawbacks. The main problem is detection of the

object when it resurfaces in the camera's frame. The appearance from the initial frame becomes inappropriate as object may change its appearance. A long-term tracker should be able to handle illumination and scale variations, partial occlusions, background clutter and operate in real-time. The idea of this project is to use tracking and learning simultaneously. A tracker provides weakly labelled training information for a detector and hence improve it during runtime. A detector can reinitialize a tracker and hence minimize the tracking failures.

In this, RAT framework divides the task of long-term tracking into three categories: Tracking, Learning, and Detection. The tracker tracks the subject in each frame. The detector detects all appearance variations observed till then and updates the tracker when essential. With the help of P-N Learning, detector's errors are estimated and the Learner updates it to avoid these mistakes henceforth. P-N learning will evaluate the detector in each frame and will estimate the errors by a pair of experts. (5)

The object can be detected and tracked in real-time and the operators can manoeuvre the camera to communicate with the object to learn the representation of the object in various circumstances. We outperform existing adaptive tracking-by-detection methods by separating object tracking and detection and hence achieve high sturdiness.

## METHODOLOGY

The task of long-time tracking is divided as follows: Tracking, Learning and Detection. (5) The tracker must follow the marked object of interest in consecutive frames. The detector restricts all observed appearances and amends the tracker when required. To evade these blunders there forth, the learning approximates the detector's blunders and re-evaluates it.

This project studies methods to recognize the detector's faults and learn from, by developing a learning method with the help of "experts" which will estimate these blunders. We call it the P-N learning.

- 1) Detections which are missed are estimated by P-expert and
- 2) False alarms by N-expert.

With the help of RAT and P-N learning, our real-time processing can be described as a highly integrated approach giving very precise object detection with RGB-D sensor with guaranteed improvements.

## *Tracking*

Tracking algorithm is used to determine the object movement. Trackers are fast, they produce smooth trajectories and require only initialization. However, during run-time, they accumulate error and fail when object disappears from the camera frame. The tracker must follow the marked object of interest in consecutive frames which is an estimation of object motion between successive frames. Trackers presume that the object of interest is perceptible throughout the sequence of frames. There are various representations of the estimation of object motion such as optical flow [12], points or contours, etc.

## *Learning*

The data tracked by a detector can be learnt and improved during run time by the learner. The tracking failures can be minimized by re-initializing it by a detector. The detector restricts all observed appearances and amends the tracker when required. To evade these blunders there forth, the learning approximates the detector's blunders and re-evaluates it.

Such a learning method should:

- (i) In intricate videos where tracking fiascos are common it must deal with it,

(ii) If the video-stream doesn't contain related info, the detector must not degrade and

(iii) Must operate in real-time.

We depend on the numerous information sources available in the video to tackle all these challenges.

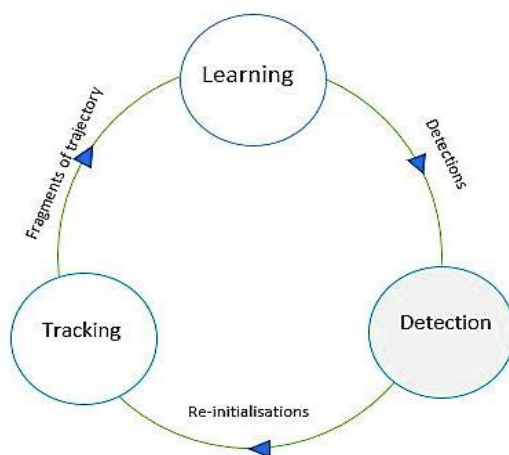
### **Detection**

The position of the subject is estimated in every frame independently by the Detection-based algorithm. Unlike Tracking, the detectors don't stray and can continue detecting if the subject moves out of camera frame. The detection localizes the subject in an input image.

### **RAT Framework**

The block diagram of the RAT framework is as shown. It is used for long-time object tracking in a sequence of frames. The components work is divided. Assuming

that the motion of object in consecutive frames is limited and it is in the camera vision. The tracker approximates the motion of the object in successive frames. Detector performs full scanning of the image by treating every frame as independent. It supports in restricting all the variations in object appearance that have been observed and learned previously. Like any other detector, this one makes two types of blunders. They are the incorrect positives and incorrect negatives. To avoid blunders hence forth, learning estimates the errors made by the detector and generates training examples. These examples are produced by assuming both the tracker and the detector may malfunction yet observes the performance of both. With the help of learning, the detector generalizes more object appearances and distinguishes against the background.



**Fig. 1** Example of qualitative results for the sequence for Face Turning



**Fig. 2** Example of distinguishing the green from red. (5)

### ***Process of RAT***

From the incoming video, we select an object of interest. The moment we select the image, it gets saved in the database which we call as the positive image database. On the next Frame, we search similar objects. The software logically selects an ROI region. At the most, the object of interest in the next frame will be within this region.

Instead of tracking the image in the whole frame, it is tracked in the Region of Interest for fast processing. In case if it is not in the Region of interest, the ROI region is increased by some amount until it reaches the full frame.

If it is detected in the region of interest, the following process executes:

We crop the ROI image from the main image and apply template matching algorithm (6). The template matching algorithm is used to search positive image in the ROI image. The template matching algorithm slides the cropped image on the target image. There are a number of template matching algorithms. All the Template matching algorithms produce different results. We use co-efficient normalised algorithm. After sliding the cropped image on the target image, it does pixel difference. It results in a grey image. The location where the grey image is the brightest is the location of the object.

The value of brightness is a measure of matching percentage. If the spot is brightest i.e., the value is 1, it is a 100% match. If it is 0, it is a 0% match. We set a threshold of matching percent, say 0.95, i.e., 95% match, we assume that the object is identified. If it is above the set threshold, the object has matched. If the object has been found, mark the object location and update screen. If it is below the set threshold say, 0.90, we set another lower threshold about 0.85 as the learning threshold. It is not exactly the same but

closer or similar to the object.

We get into learning stage.

Where this image has been compared to the image set by the user, in the positive image database and has produced the value lower than the matching threshold, we now compare and do template matching with the background. The background is divided into grids and each grid is known as negative image. We create an array of n images. This image is compared to n images and if there is a match, then we add it to the n image array i.e., background and is rejected.

The features keep on changing and the array becomes full. We can set a limit on the array, say 50 pictures. The following process takes place until the array becomes full. When the array becomes full, the newly learned image is added as the latest image and the image with the lowest matching value is taken out.

### **RESULT**

1. The PCA corresponding proportion up to 0.98
2. Average localization error is condensed up to 65 %.
3. Template matching intensity is in the range of 0-1 where 0 is dark and 1 is bright.
4. Development in object learning is due to the adaptation of PN learning Algorithm
5. This technique can track the moving object positively; furthermore, the trace of object can be also extracted and the tracking result is undisturbed with the proportions of moving object. This technique still can track the interested target and filter out the noise by examining the size of moving regions
6. The speed of system for processing a frame in the ROI is about 300 ms approximately and declines if ROI grows.

7. If ROI miscarries for fast motion video then Background deduction procedure is utilised.

### **Test Cases**

Steps for Execution

### **Initialization Procedure**

1. Select object of interest: Operator will select the object to be tracked from the

incoming live video stream.

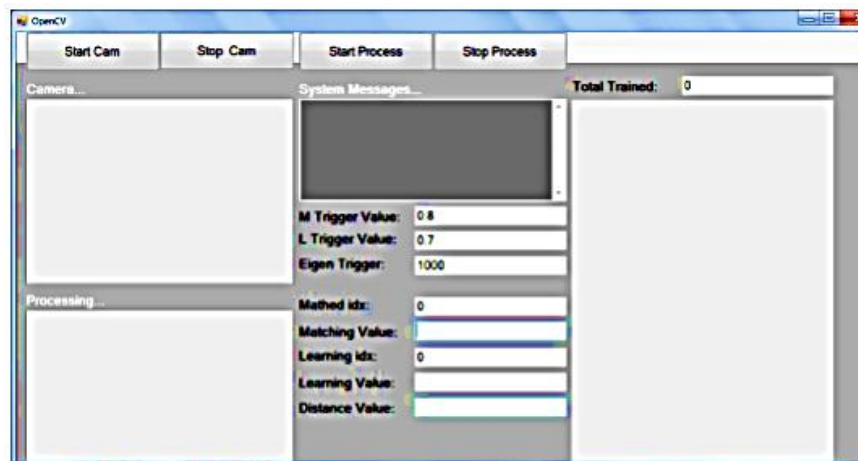
2. Load the image in the database: Selected or cropped image will be saved in positive objects database.
3. ROI Generation: The system will automatically create the region of interest (ROI) at 20 pixels area more than the selected position in the frame and will search the object in the ROI.

### **Procedure for Detection**

Test Case Description	Test Input	Expected Output	Actual Output
To check valid graphical user interface (GUI) on screen	Display of main functions to be performed	Home Page Should be displayed	Main GUI Display window should be displayed
To check camera initialization process	User should click button of Start camera input	System should display the image in the camera window	Image is displayed in the camera window.
Checking of object selection in the image in the camera window	User should click button of start process	Message should be displayed "processing stopped" Please select the object	Message is displayed "Processing stopped" Please select the object
To check whether region of interest (ROI) of cropped image is displayed or not	Selection of Cropped image from user	Cropped image should be displayed in the processing box	Image is displayed in the processing box
To check trained images	Selection of cropped image from the user.	Cropped object should be displayed in the trained images box.	Cropped object is displayed in total trained box along with count.

1. Get the live frame
2. By template matching algorithm, search for specific index image from the trained object database and the cropped image.
3. Grow the ROI if there is no match found and reiterate above step.
4. Mark the new position object {update the tracker } if match found. Matching percentage feature is found by applying eigen objector algorithm {PCA} principle
5. component analysis.
6. If matching percentage is less than the set matching limit but above the learning percentage, do template matching for background. If it matches any of the background, add it in the negative database. If no match found in the negative database, then learn the object by storing its new image in [positive image database.
7. Reiterate above steps.





*Fig. 3 Executing screen*

Above figure is valid graphical user interface on the screen when in execution.



*Fig. 4 Test image*

Name of object of interest: Brown Object

Frames	Moving Camera	Partial Occlusion	Full Occlusion	Pose and illumination Change
Yes	Yes	Yes	Yes	Yes

## CONCLUSION

In this project, an enhanced RAT is built. By equating with the prevailing target detection of an unidentified object in a video stream approaches, RAT is effective and perceives the objects faster. Any changes in appearance of the target is adjusted and it does not falter when the target moves in and out of the camera view. It decomposes the task into three components: tracking, learning and detection. The learning segment examination demonstrates that a target detector can be trained from a single

example and an unlabelled video stream using the following strategy:

1. evaluate the detector,
2. estimate its blunders by a couple of experts, and
3. apprise the classifier.

By developing this cohesive RAT structure with RGB-D data, we can operate them easily. As compared to current practices which require severely supervised laboratory environment, this method operates at low cost. Positively tracked templates are used as p-type training information. The target object can thus

provide online learning. The depth data provides features like surface normal and colour gradient which help in precise detection of the object of interest in tough circumstances. Each expert focuses on identifying a specific form of the classifier blunder. Although they themselves are permitted to make blunders. The constancy of the learning is attained by designing experts that mutually compensate their blunders. The formalization of this process as a discrete dynamical system is the theoretical contribution of this project. It allows specifying conditions under which the learning process assures improvement of the classifier. Space-Time relationships can be exploited by these experts in the video. RAT framework is a real-time approach.

The new algorithm can automatically accomplish RAT initialization and has better adaptation in cases of full out-of-plane rotation and strong deformation. Experiments have shown the stability, robustness and better performance than existing methods

Most importantly, the trained object detector can be exploited in many real-time applications in real world.

### Applications

1. Recognition based on motion.
2. Automated detection of scenes to detect activities.
3. Gesture recognition for human-computer interaction.
4. To gather real time traffic statistics.
5. To develop navigation of vehicle which can be used for video-based path planning and avoidance obstacle.
6. Security
7. Domestic services
8. Sports competition

### REFERENCE

1. C. H. Lampert, M. B. Blaschko, and T. Hofmann. Beyond sliding

windows: Object localization by efficient subwindow search. In IEEE Conference on Computer Vision and Pattern Recognition, pages 1–8, 2008.

2. S.-C. S. Cheung and C. Kamath. Robust techniques for background subtraction in urban traffic video. In Visual Communications and Image Processing 2004 (Proceedings Volume), volume 5308, pages 881–892. SPIE, 2004.
3. R. J. Radke, S. Andra, O. Al-Kofahi, and B. Roysam. Image change detection algorithms: a systematic survey. IEEE Transactions on Image Processing, 14(3):294–307, Mar. 2005.
4. P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features.
4. In 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, volume 1, pages I-511–I-518, Los Alamitos, CA, USA, Apr. 2001. IEEE Comput. Soc.
5. Z. Kalal, J. Matas, and K. Mikolajczyk, “P-N Learning: Bootstrapping Binary Classifiers by Structural Constraints,” Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2010.
6. Open CV documentation
7. Marwa abdel el Azeem Marzouk, —Modified background subtraction algorithm for motion detection in surveillance systems, vol 1, Number 2, (2010), pp -112123.
8. Priti Kuralkar\*, Prof. V.T.Gaikwad —Background Subtraction and Shadow Detection Techniques A Review Paper, International Journal of Computer, Electronics & Electrical Engineering (ISSN: 2249 –9997) Volume 2– Issue1
9. The OpenCV Tutorials, Release 2.4.2 July, page 1-355

10. JainHong, Zhang, Zhenuan, Wei Guo, —Object tracking using improved camshaft with SURF method, OSSC2011 May-2012, 987-61284-495-4/11 @2011IEEE
11. R. Collins, Y. Liu, and M. Leordeanu, “Online Selection of Discriminative Tracking Features,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, no. 10, pp. 1631–1643, 2005.
12. L. Wang, W. Hu, and T. Tan, “Recent developments in human motion analysis,” Pattern Recognition, vol. 36, no. 3, pp. 585–601, 2003